because of inadequate genealogical documentation: one because of nonpaternity (mother was the affected parent) mentioned above and another because we could not document the reported grandparental relationship. A recombinant found in a distant branch of family Z (shown at the left end of Fig. 1) was not included because of the complexity of computer entry of such a distant relative; she would have contributed little to the recombination fraction either way.

Our results confirm linkage between the HD locus and the G8 probe in two additional families. These families, both with pathologically confirmed HD cases, differ in ethnicity, age-at-onset distribution, presence of juvenile cases, neurological manifestations, and presence and type of psychiatric disorder. Finding linkage to G8 in these families supports the hypothesis that all families with the HD phenotype have a mutation at the same chromosome-4 locus linked to G8. In fact, even though our estimate of the recombination fraction between the HD locus and the G8 probe is greater than that previously reported (*6*), the 95 percent confidence intervals of the two sets of data overlap (0 to 12 percent as compared to 0 to 6 percent). Furthermore, our findings suggest that phenotypic differences seen in the two families may be due either to allelic differences or to influences from unlinked loci.

By means of three endonucleases that detect four or five polymorphic restriction sites, the proportion of family members who could be assigned unambiguous haplotypes increased, thus improving the utility of the linkage analysis (*18*). This increase in the proportion of persons with unambiguous haplotype assignments accounts for some of the difference between the estimate of the recombination fraction for families D and Z and that in the original report (*6*).

If members of these two families had undergone genetic counseling with the G8 probe results, there would have been a 12 percent error rate in counseling of genotype assignment (Table 2). This finding and the remaining uncertainty about the universality of linkage in all HD families (this report brings the number of reported kindreds to only four) suggests that further studies are needed before applying G8 linkage studies to clinical practice.

References and Notes

1. G. Huntington, *Med. Surg. Rep. (Philadelphia)* **26**, 317 (1872).
2. G. W. Bruyn, in *Handbook of Clinical Neurology*, P. J. Vinken and G. W. Bruyn, Eds. (North-Holland, Amsterdam, 1968), vol. 5, pp. 298–378.
3. S. E. Folstein and M. F. Folstein, *Psychiatr. Dev.* **2**, 193 (1983).
4. A. D. Merritt *et al.*, in *Progress in Neurogenetics*, A. Barbeau and J. R. Brunette, Eds. (Excerpta Medica, Amsterdam, 1969), pp. 645–650.
5. A. Tyler and P. S. Harper, *J. Med. Genet.* **20**, 179 (1983).
6. J. F. Gusella *et al.*, *Nature (London)* **306**, 234 (1983).
7. R. M. Lawn *et al.*, *Cell* **15**, 1157 (1978).
8. J. F. Gusella *et al.*, *Science* **225**, 1320 (1984).
9. S. E. Folstein, M. H. Abbott, M. L. Franz, *J. Neurogenet.* **1**, 175 (1984).
10. S. E. Folstein *et al.*, *Neurobehav. Toxicol. Tetratol.* **5**, 605 (1983).
11. S. E. Folstein *et al.*, *Psychol. Med.* **13**, 537 (1983).
12. S. H. Orkin *et al.*, *N. Engl. J. Med.* **299**, 166 (1978).
13. J. F. Gusella *et al.*, in preparation.
14. J. A. Phillips *et al.*, *J. Clin. Invest.* **70**, 489 (1982).
15. A. R. Wyman and R. White, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6754 (1980).
16. J. Ott, *Am. J. Hum. Genet.* **26**, 588 (1974).
17. M. A. Pericak-Vance *et al.*, *Am. J. Med. Genet.* **14**, 49 (1983).
18. J. Gusella *et al.*, *Am. J. Hum. Genet.* **36** (Suppl.), 1395 (1984).

19. Supported by grants NS 16375, AM 13983, and NS 16367 from the National Institutes of Health and by grants from the Hereditary Disease Foundation and the Julieanne Dorn Fund for Neurological Research.

22 January 1985; accepted 11 June 1985

Diversity of Circumsporozoite Antigen Genes from Two Strains of the Malarial Parasite *Plasmodium knowlesi*

Abstract. *The complete nucleotide sequence of the coding region of the circumsporozoite antigen gene (CS gene) of the Nuri strain of the malarial parasite* Plasmodium knowlesi *is presented. The gene from the Nuri strain exhibits a novel form of sequence diversity when compared to the CS gene from the H strain. Instead of the 12 tandem repeating 36–base pair units of the H strain, the Nuri strain contains 16 tandem repeating 27–base pair units of a different nucleotide sequence that encodes a different repeating peptide. In contrast, the 5′ and 3′ coding and noncoding sequences flanking the repeats are 98 percent conserved in both strains.*

SHOBHONA SHARMA*
PAMELA SVEC
Department of Biochemistry, New York University Medical Center, New York 10016
G. H. MITCHELL
Department of Chemical Pathology, Guy's Hospital Medical School, London SE1 9RT, England
G. NIGEL GODSON
Department of Biochemistry, New York University Medical Center

*Present address: Departments of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710

The gene for the circumsporozoite antigen (CS gene) of the H strain of the simian malarial parasite *Plasmodium knowlesi* has recently been cloned and sequenced (*1*). A remarkable feature of this gene is the presence of 12 tandemly repeating 36–base pair (bp) units in the middle of the coding sequence. This gene encodes a parasite surface protein with approximately 40 percent of the polypeptide chain consisting of a tandemly repeating 12–amino acid peptide. Since then, a number of surface antigens from the sporozoite and merozoite stages of different malarial parasites (*P. falciparum*, *P. cynomolgi*, and *P. lophurae*) have been cloned, and each of the surface antigen genes has been found to contain tandemly repeating units (*2–7*). We now present the complete nucleotide sequence of the CS gene of another *P. knowlesi* strain, the Nuri strain (Fig. 1). The gene from the Nuri strain exhibits a different nucleotide repeat unit than that from the H-strain gene; the 5′ and 3′ coding and noncoding sequences flanking the repeat units are the same in both strains. Conserved sequences flanking variable repeat regions have recently been reported for S-antigen genes from two isolates from *P. falciparum* (*8*).

To isolate the Nuri strain CS gene, we constructed a λgt11 (*9*) library with blood-form DNA randomly cleaved with DNase I to an average size of 1000 bp (Fig. 2). Using as a probe a 1.6–kilobase pair (kb), Aha III-restriction endonuclease fragment that contained the CS gene of the H strain (*1*) and was labeled with ^{32}P, we isolated and plaque-purified 11 positive Nuri strain clones from approximately 120,000 plaques. A restriction-endonuclease cleavage map of the Nuri strain clones was constructed and compared with that of the H strain DNA (Fig. 2). The cleavage sites appeared to be identical for 2.5 kb in both strains. Four λgt11 clones, λKN5, λKN6, λKN7, and λKN8, were used to sequence the Nuri strain CS gene. The nucleotide sequence of the CS gene was established by the Sanger dideoxy chain–termination method (*10*) after the inserts were subcloned into M13 sequencing vectors (*11*).

The complete nucleotide sequence and the deduced amino acid sequence of the Nuri strain CS gene is shown in Fig. 2. For the regions 5′ and 3′ to the repeating units, the sequence in the Nuri strain is 98 percent identical to that in the H strain. The diversity of the sequence of repeating units, however, is striking.

There are 14 perfect and 2 partial repeat units of 27 bp each in the Nuri strain compared to 12 repeat units of 36 bp each in the H strain. Both partial repeat units are on the 3′ end of the repeat region and none are at the 5′ end (Fig. 1). Only one base-pair change from G to A was observed in three of the first 14 repeat units in the Nuri strain, and this results in a change in the coded amino acid from glycine to arginine. This change from an uncharged amino acid to a charged basic residue is likely to cause a conformational change that might lead to an antigenically distinct epitope for these three nonapeptide units compared to the rest of the 11 repeats. A similar change of a glycyl residue to an aspartyl residue has recently been reported in the repeat units of the CS protein of *P. cynomolgi* (*7*). In contrast, all the changes observed in the repeat units of the CS gene of the H strain are in the third position of the codon, which causes no change in the composition of the repeat peptide composition (*1*). One peculiar feature of the Nuri strain repeat units is the total lack of T residues. The repeat sequences of the CS genes in other malarial parasites also contain a low proportion of T residues [8 percent in *P. falciparum* (*3*) and the H strain of *P. knowlesi* (*1*) and 16 percent in *P. cynomolgi* (*7*)]. This is in contrast to the overall *Plasmodium* DNA that is rich in A and T residues (approximately 80 percent).

The CS genes of the H and Nuri strains have some common features. Each repeat unit contains one proline residue and one acidic residue (a glutamyl residue and an aspartyl residue in the Nuri and H strain, respectively). There are also three glycine and three alanine residues in each repeat unit in both the strains. However, although the composition is similar, the sequence of the amino acids of the repeat peptides in the two strains bear no significant resemblance (Fig. 3).

In contrast to the repeating units of the gene, only minor nucleotide changes were observed in the 5′ and 3′ flanking sequences. The nucleotide (and amino

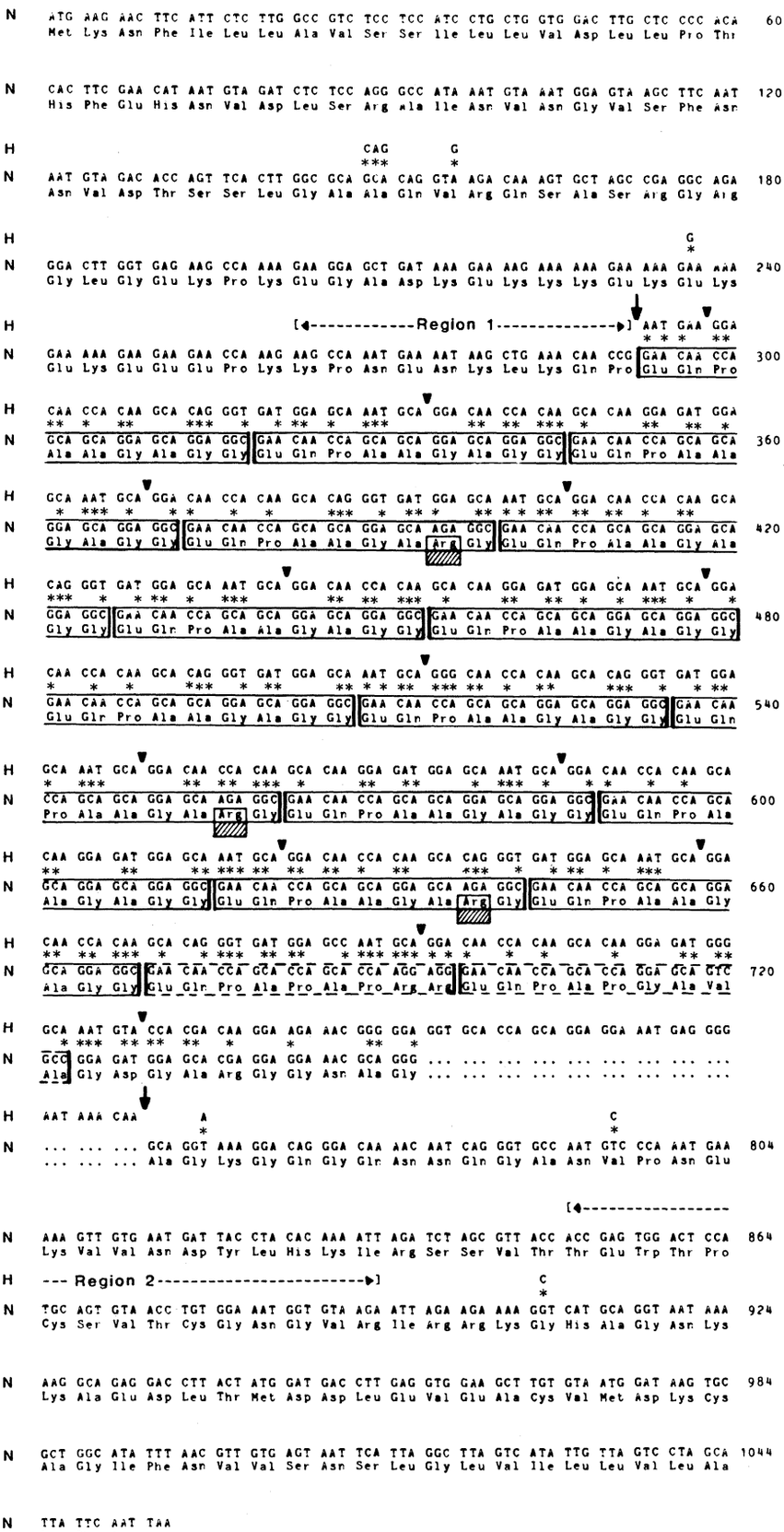

Fig. 1. Complete nucleotide sequence of the coding region and the deduced amino acid sequence of the CS gene of the Nuri strain of *P. knowlesi*. Only the complete nucleotide sequence of the repeat region of the H strain is identified by arrows. On the 5′ and 3′ sides of the repeat region, the sequences of the H and Nuri strains are almost identical, and therefore only the mismatches in the nucleotide sequence (*) are shown for the H strain. Boxed regions are the Nuri repeating units; (▼) indicates the junctions of the repeats of the H strain. Broken boxes represent the last two partial repeats of the Nuri strain. The base changes that occur in the 4th, 10th, and 13th of the Nuri repeat units, are shown by the hatched boxes. The two regions of the protein that are nearly identical to regions in the CS protein of *P. falciparum* (*3*) are marked region 1 and region 2. The clones used for obtaining the CS gene sequence of the Nuri strain are shown in Fig. 2.

acid) sequence from the ATG initiation codon of the CS gene coding region (nucleotide 1, Fig. 1) to the start of the tandem repeating units (nucleotide 291) is essentially identical to both strains with only five nucleotide changes, which result in two amino acid changes out of 97 amino acids. On the carboxyl terminal side of the tandem repeating units, starting from base pair 754, the nucleotide and amino acid sequences are again almost completely homologous with three nucleotide differences out of 300, which result in two amino acid changes out of 100 amino acids. The 5′ and 3′ noncoding regions of the gene have been sequenced for 400 and 200 bp, respectively, and are nearly identical.

Two small sections in the CS genes of *P. falciparum* and the H strain of *P. knowlesi* are homologous (*3*). These sections are also maintained in the Nuri strain of *P. knowlesi* (Fig. 3). The diversity of the Nuri strain gene begins just beyond the homologous section in region 1. The conservation of these sections in the Nuri strain gene supports the possibility that they encode important functions.

The CS protein of the Nuri strain is 12 amino acids shorter than that of the H strain. However, in terms of overall protein structure, the Nuri protein contains the same signal sequence, the same charged region before the repeats, and the same anchor region as in the CS protein of the H strain. The major differences between the proteins are the size, number, and composition of the peptide repeat units of these strains.

Southern blot analysis indicated an allelic nature for the CS genes in these two strains. Blots of Eco RI and Aha III restriction enzyme digests of genomic DNA from the blood form of the Nuri strain and the H strain showed similar single bands when probed with a 1.6-kb Aha III fragment containing the CS gene from the H-strain. The presence of a single band is consistent with the CS gene being a single gene in the Nuri strain, as it is in the H strain (*1*).

The production of a family of genes with a strain-specific variation in a tandemly repeated internal section flanked by a constant region requires two molecular mechanisms. First, a single unit of the repeat section must be amplified; second, the repeat region must be able to behave independent of the rest of the gene. Whether these mechanisms operate within the life cycle of *Plasmodium*, rather than on an evolutionary time scale, remains to be determined. There is, however, no difference between the CS gene copies of the merozoite and sporozoite DNA, but this does not rule out changes in other developmental stages such as the gametocyte or zygote. The precise interposition of the changed repeat sequence at the 5′ end of the CS gene and the imprecise interposition at the 3′ end (Fig. 2) probably reflect the underlying molecular mechanisms.

The presence of repeating peptide units in all malarial surface antigens has given rise to speculation regarding their function (*2, 12*). We have suggested that they have a role in evasion of the host immune system by acting as an immune decoy protein (*12*). It has also been suggested that the presence of repeating peptide units creates an immunodominant region (*7*). Repeating peptide units alone are clearly insufficient because they occur in proteins such as silk and collagen that are not highly immunogenic (*13, 14*). At some point in its evolution, the malarial parasite may have generated a mechanism for amplification of already existing immunogenic epitopes, giving rise to tandem repeating units. Parasites containing such antigens may then have been selected because of their ability to evade the immune system.

Of all the surface antigens of the malarial parasite, the CS antigen has been considered a suitable candidate for a

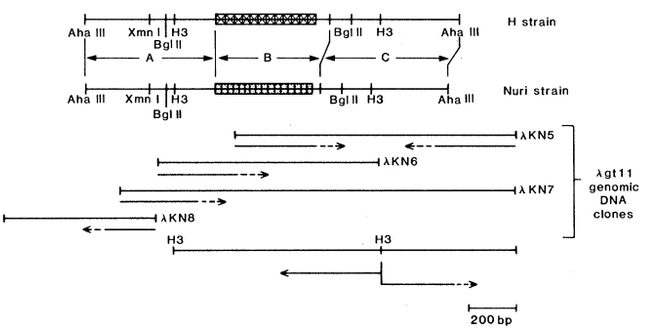

Fig. 2. Restriction maps of genomic DNA containing the CS gene of the H strain (*1*) and the Nuri strain of *P. knowlesi*. The clones used for restriction endonuclease analysis and sequencing were obtained from a λgt11 library constructed from a partial DNase-I digest of genomic DNA from the blood form of the Nuri strain isolated, as described (*21*), from erythrocytes of rhesus monkeys infected with *P. knowlesi*. Parasite DNA (10 μg) was digested with DNase I in the presence of $MnCl_2$ under optimum conditions for generating 0.5- to 1.5-kb fragments (*22*). These were fractionated on a 5 to 20 percent sucrose gradient by centrifugation at 35,000 rev/min for 14 hours (Beckman SW41 rotor). Fractions containing 0.5- to 1.5-kb fragments were pooled, and the DNA was precipitated. This DNA (2 μg) was methylated with Eco RI methylase and then ligated with Eco RI linkers. Approximately 200 ng of the linked DNA fragments were ligated to λgt11 Eco RI arms (*9*), previously treated with calf intestine alkaline phosphatase, and packaged in vitro (Boehringer Mannheim). Recombinants (1.6×10^6) with less than 5 percent background nonrecombinant plaques from λgt11 were obtained with *Escherichia coli* strain Y1088 (*23*). Plaques (12×10^4) were screened with ^{32}P-labeled nick-translated Aha III fragment containing the CS gene of H strain (*1*) by the method of Benton and Davis (*24*). The arrows show the origin, direction, and extent of the sequences. The boxed bars indicate the repeating nucleotide sequences in the two strains. B represents the repeat regions; A and C represent the 5′ and 3′ sequences flanking the repeat units.

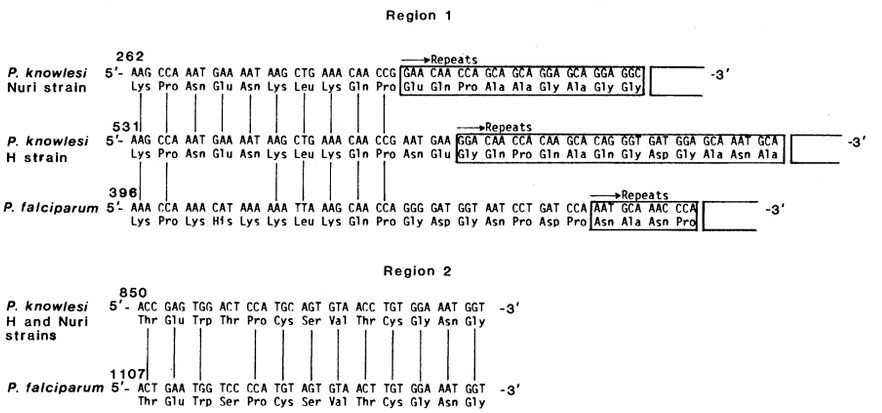

Fig. 3. Regions of identity between the CS genes of *P. knowlesi* and *P. falciparum* (*3*). The location of this sequence is indicated by the nucleotide numbers from Fig. 1 for the Nuri strain and from (*1*) and (*3*) for the H strain of *P. knowlesi* and *P. falciparum*, respectively. In the case of region 2, the nucleotide sequence for *P. knowlesi* is identical in the H strain and the Nuri strain. The nucleotide numbers at the start of region 2 are 1167 for the H strain (*1*) and 850 for the Nuri strain; only the latter number is shown. The matching amino acids are shown by vertical lines. The boxed sequences indicate the repeat units.

vaccine (*15*) because of the strain independence of the CS proteins. Extensive cross-reactivity has been observed between the CS proteins from different strains of *P. cynomolgi* (*16*), *P. vivax* (*17*), and *P. falciparum* (*17*). In the case of *P. knowlesi*, only two strains (H and P strains) have been tested, and they showed weak cross-reactivity (*15*). Since the repeating units constitute the immunogenic epitope, recent emphasis has been on making protective antigens by means of synthetic peptide units coupled to a carrier molecule (*18*). In view of the extreme differences in the repeating peptide units in the two *P. knowlesi* strains, the utility of such synthetic protective antigens may be limited. Indeed, the CS protein of the Nuri strain does not cross-react with the antibody to the repeat region of the H strain CS protein (*19*). Recent reports (*7, 20*) have indicated possible diverse repeating units and serological antigenic diversity between the CS proteins of different strains of another simian malarial parasite, *P. cynomolgi*. It remains to be determined whether such differences occur in the CS proteins of different strains of human malarial parasites.

References and Notes

1. L. S. Ozaki, P. Svec, R. S. Nussenzweig, V. Nussenzweig, G. N. Godson, *Cell* **34**, 815 (1983).
2. R. L. Coppel *et al.*, *Nature (London)* **306**, 751 (1983).
3. J. B. Dame *et al.*, *Science* **225**, 593 (1984).
4. R. L. Coppel *et al.*, *Nature (London)* **310**, 789 (1984).
5. M. Koenen *et al.*, *ibid.* **311**, 382 (1984).
6. J. V. Ravetch, R. Feder, A. Pavlovec, G. Blobel, *ibid.* **312**, 616 (1984).
7. V. Enea *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7520 (1984).
8. A. F. Cowman *et al.*, *Cell* **40**, 775 (1985).
9. R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194 (1983).
10. F. Sanger, S. Nicklen, R. Coulson, *ibid.* **74**, 5463 (1977).
11. J. Messing and J. Vieira, *Gene (Amsterdam)* **19**, 269 (1982).
12. G. N. Godson, J. Ellis, J. R. Lupski, L. S. Ozaki, P. Svec, *Philos. Trans. R. Soc. London Ser. B* **307**, 129 (1984).
13. R. F. Doolittle, in *The Evolution of Proteins*, H. Neurath and R. L. Hill, Eds. (Academic Press, New York, 1979), vol. 4, pp. 1–118.
14. C. Steffen and R. Timpl, *Int. Arch. Allergy Appl. Immunol.* **22**, 333 (1963).
15. A. H. Cochrane, R. S. Nussenzweig, E. H. Nardin, in *Malaria*, J. P. Kreier, Ed. (Academic Press, New York, 1980), vol. 3, pp. 163–202.
16. D. H. Chen, R. S. Nussenzweig, W. E. Collins, *J. Parasitol.* **62**, 636 (1976).
17. E. H. Nardin *et al.*, *J. Exp. Med.* **156**, 20 (1982).
18. J. Gysin, J. Barnwell, D. H. Schlesinger, V. Nussenzweig, R. S. Nussenzweig, *ibid.* **160**, 935 (1984).
19. S. Sharma, unpublished results.
20. A. H. Cochrane *et al.*, *Mol. Biochem. Parasitol.* **14**, 111 (1985).
21. M. Goman *et al.*, *ibid.* **5**, 391 (1982).
22. S. Anderson, *Nucleic Acids Res.* **9**, 3015 (1981).
23. R. A. Young and R. W. Davis, *Science* **222**, 778 (1983).
24. W. D. Benton and R. W. Davis, *ibid.* **196**, 180 (1977).

25. We thank F. Rheinhart and L. S. Ozaki for preparing the DNA and S.-C. Chiang for assistance in making the λgt11 library. Supported by the National Institutes of Allergy and Infectious Diseases grant 1 R01 AI21496-01.

4 March 1985; accepted 29 May 1985

Selective Attention Gates Visual Processing in the Extrastriate Cortex

Abstract. *Single cells were recorded in the visual cortex of monkeys trained to attend to stimuli at one location in the visual field and ignore stimuli at another. When both locations were within the receptive field of a cell in prestriate area V4 or the inferior temporal cortex, the response to the unattended stimulus was dramatically reduced. Cells in the striate cortex were unaffected by attention. The filtering of irrelevant information from the receptive fields of extrastriate neurons may underlie the ability to identify and remember the properties of a particular object out of the many that may be represented on the retina.*

JEFFREY MORAN
ROBERT DESIMONE
Laboratory of Neuropsychology, Building 9, Room 1N107, National Institute of Mental Health, Bethesda, Maryland 20205

Our retinas are constantly stimulated by a welter of shapes, colors, and textures. Since we are aware of only a small amount of this information at any one moment, most of it must be filtered out centrally. This filtering cannot easily be explained by the known properties of the visual system. In primates, the visual recognition of objects depends on the transmission of information from the striate cortex (V1) through prestriate areas into the inferior temporal (IT) cortex (*1*). At each successive stage along this pathway there is an increase in the size of the receptive fields; that is, neurons respond to stimuli throughout an increasingly large portion of the visual field. Within these large receptive fields will typically fall several different stimuli. Thus, paradoxically, more rather than less information appears to be processed by single neurons at each successive stage. How, then, does the visual system limit processing of unwanted stimuli? The results of our recording experiments on single neurons in the visual cortex of trained monkeys indicate that unwanted information is filtered from the receptive fields of neurons in the extrastriate cortex as a result of selective attention.

The general strategy of the experiment was as follows. After isolating a cell, we first determined its receptive field while the monkey fixated on a small target. On the basis of the cell's response to bars of various colors, orientations, and sizes, we determined which stimuli were effective in driving the cell and which were ineffective. Effective stimuli were then presented at one location in the receptive field concurrently with ineffective stimuli at a second location. The monkey was trained on a task that required it to

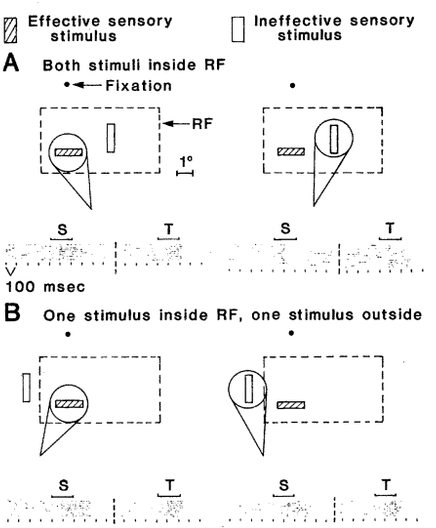

Fig. 1. Effect of selective attention on the responses of a neuron in prestriate area V4. (A) Responses when the monkey attended to one location inside the receptive field (RF) and ignored another. At the attended location (circled), two stimuli (sample and test) were presented sequentially and the monkey responded differently depending on whether they were the same or different. Irrelevant stimuli were presented simultaneously with the sample and test but at a separate location in the receptive field. In the initial mapping of the receptive field, the cell responded well to horizontal and vertical red bars placed anywhere in the receptive field but not at all to green bars of any orientation. Horizontal or vertical red bars (effective sensory stimuli) were then placed at one location in the field and horizontal or vertical green bars (ineffective stimuli) at another. The responses shown are to horizontal red and vertical bars but are representative of the responses to the other stimulus pairings. When the animal attended to the location of the effective stimulus at the time of presentation of either the sample (S) or the test (T), the cell gave a good response (left), but when the animal attended to the location of the ineffective stimulus, the cell gave almost no response (right), even though the effective stimulus was present in its receptive field. Thus the responses of the cell were determined by the attended stimulus. Because of the random delay between the sample and test stimulus presentations, the rasters were synchronized separately at the onsets of the sample and test stimuli (indicated by the vertical dashed lines). (B) Same stimuli as in (A), but the ineffective stimulus was placed outside the receptive field. The neuron responded similarly to the effective sensory stimulus, regardless of which location was attended.